

## Ficha de Unidade Curricular – (Versão A3ES 2018-2023)

### 1. Caracterização da Unidade Curricular.

1.1. **Designação da unidade curricular** (1.000 carateres).  
Processamento de Fala / Speech Processing

1.2. **Sigla da área científica em que se insere** (100 carateres).  
INF

1.3. **Duração**<sup>1</sup> (100 carateres).  
Semestral

1.4. **Horas de trabalho**<sup>2</sup> (100 carateres).  
162

1.5. **Horas de contacto**<sup>3</sup> (100 carateres).  
67,5

1.6. **ECTS** (100 carateres).  
6

1.7. **Observações**<sup>4</sup> (1.000 carateres).  
Optativa

1.7. **Remarks** (1.000 carateres).  
Optional

2. **Docente responsável e respetiva carga letiva na Unidade Curricular** (preencher o nome completo) (1.000 carateres).  
Carlos Eduardo de Meneses Ribeiro, 67,5 horas de contacto

3. **Outros docentes e respetivas cargas letivas na unidade curricular** (1.000 carateres).  
Não se aplica

4. **Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes)**. (1.000 carateres).

Os seres humanos há muito que deixaram de ser o único gerador ou destinatário da fala, podendo parte desta cadeia ser implementada por sistemas automáticos. Estes sistemas encontram-se em PC, smartphones, acesso a BD, etc. O objeto do processamento de sinais de fala é a representação digital dos sinais de fala, análise e extração de características e o desenvolvimento de modelos de síntese e reconhecimento. Todas estas ferramentas são cruciais na implementação de sistemas de comunicação falada, seja esta comunicação à distância ou comunicação homem-máquina.

Os estudantes que terminam com sucesso esta unidade curricular serão capazes de:

1. Compreender a estrutura do sinal de fala;
2. Compreender as principais tecnologias de fala: codificação, síntese e reconhecimento;
3. Desenvolver e avaliar projetos na área de codificação e reconhecimento;
4. Escrever relatórios onde se descrevem os projetos implementados, se analisam e comparam resultados e se justificam decisões.

4. **Intended learning outcomes (knowledge, skills and competences to be developed by the students)**. (1.000 characters).

Humans have long ceased to be the sole generator or destination of speech, and part of this chain may be implemented by automatic systems. These systems are in PC, smartphones, access to DB, etc. The object of speech signal processing is the digital representation of speech signals, characteristic analysis and extraction, and the development of synthesis and recognition models. All these tools are crucial in the implementation of spoken communication systems, be it distance communication or man-machine communication.

A student completing this course unit should be able to:

1. Understand the speech signal structure;

2. Understand the main speech technologies: coding, synthesis and recognition;
3. Develop coding and recognition projects;
4. Write reports describing the implemented projects and where results are analyzed and compared and decisions are justified.

**5. Conteúdos programáticos (1.000 caracteres).**

- I. Produção e modelação de fala. Classificação fonética.
- II. Análise de fala: predição linear, estimação de coeficientes LSF e de formantes, análise cepstral, estimação da frequência fundamental.
- III: Síntese de fala: predição linear, síntese sinusoidal e de formantes.
- IV. Atributos dos codificadores. Métodos de codificação de fala: vocoders, análise por síntese, codificação sinusoidal, MELP, codificação a muito baixo débito binário.
- V. Reconhecimento de fala e de orador. Características para reconhecimento de fala e de orador. Métodos de reconhecimento: DTW e HMM.

**5. Syllabus (1.000 characters).**

- I. Speech production and modelling. Phonetic classification.
- II. Speech analysis: linear prediction, LSF coefficients and formants estimation, cepstral analysis, fundamental frequency estimation.
- III. Speech synthesis: linear prediction, sinusoidal and formant synthesis.
- IV. Coders attributes. Speech coding methods: vocoders, analysis by synthesis, sinusoidal coding, MELP, coding at very low bit rate.
- V. Speech and speaker recognition. Features for speech and speaker recognition. Recognition methods: DTW and HMM.

**6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular (1.000 caracteres).**

O objetivo de aprendizagem 1 é atingido através da lecionação dos conteúdos programáticos I e II. O objetivo de aprendizagem 2 é atingido através da lecionação dos conteúdos programáticos III e IV e V. O objetivo de aprendizagem 3 é atingido através do desenvolvimento de pequenos projetos relativos aos conteúdos programáticos dos conteúdos programáticos II, IV e V. O objetivo de aprendizagem 4 é atingido através da escrita e discussão dos relatórios dos projetos realizados.

**6. Evidence of the syllabus coherence with the curricular unit's intended learning outcomes (1.000 characters).**

Learning outcome 1 is achieved through the teaching of program contents I and II. Learning outcome 2 is achieved by teaching program contents III and IV and V. Learning outcome 3 is achieved through the development of small projects related to the program content of program contents II, IV and V. The learning outcome 4 is achieved through writing and discussing project reports.

**7. Metodologias de ensino (avaliação incluída) (1.000 caracteres).**

A metodologia de ensino desenvolvesse em várias componentes:

T – 30 horas de contacto teóricas - Exposição e discussão dos conceitos teóricos, incentivando à interatividade e colocação de questões;

TP – 15 horas teórico-práticas: São resolvidos exercícios exemplificativos da matéria teórica;

PL – 22,5 horas de contato de prática laboratorial: Os conceitos teóricos são aprofundados através da implementação de projetos, realizados em grupo.

Todos os objetivos são avaliados através do acompanhamento dos projetos (40%), qualidade dos relatórios (20%) e discussão final (40%).

**7. Teaching methodologies (including assessment) (1.000 characters).**

The teaching methodology is developed in several components:

T – 30 theoretical teaching contact hours - Presentation and discussion of theoretical concepts, interactivity and asking questions are encouraged;

TP - 15 theoretical and practical teaching contact hours: Exemplary exercises are solved;

PL - 22.5 laboratory practice contact hours: Theoretical concepts are further developed through the implementation of projects, performed in groups.

All learning outcomes are evaluated through the project monitoring (40%), report quality (20%) and final discussion (40%).

**8. Demonstração da coerência das metodologias de ensino com os objetivos de aprendizagem da unidade curricular (3.000 carateres).**

Nas aulas teóricas são expostos os conceitos relativos aos conteúdos programáticos a que correspondem os objetivos de aprendizagem – É facultado um conjunto de perguntas teóricas como motivação para o estudo fora das horas de contacto.

Nas aulas teórico-práticas reforçam-se os objetivos de aprendizagem.

Nas aulas de prática laboratorial pretende-se que os estudantes desenvolvam projetos que são a base da sua aprendizagem.

Na discussão final é discutido e o conteúdo do relatório dos projetos, tendo em consideração de um modo geral todos os objetivos de aprendizagem.

**8. Evidence of the teaching methodologies coherence with the curricular unit's intended learning outcomes (3.000 characters).**

In the lectures are exposed the syllabus corresponding to the learning objectives. It is provided with a set of theoretical questions as motivation to study outside the contact hours.

The practical classes reinforce to the learning objectives.

In the laboratory practical classes is intended that students develop projects that are the basis of their learning.

In the final discussion is discussed the projects contents, taking in consideration all learning objectives.

**9. Bibliografia de consulta/existência obrigatória (1.000 carateres).**

Carlos Meneses, Processamento Digital de Fala, Publicação pessoal, 2000-2012

[Delgado-Martins (88)] - M. R. D. Martins, "Ouvir Falar - Introdução à Fonética Portuguesa", Editorial Caminho, Coleção Universitária, Série Linguística, 1989.

[Trancoso (90)] - I. M. Trancoso, J. S. Marques, C. M. Ribeiro, "Two Solutions for Speech Coding at 4.8-9.6 KBPS", Speech Communications Journal, Vol 9 5/6, pp.389-400, Dezembro de 1990.

[Soong (84)] - F. Soong, B. Juang, "Line Spectrum Pair (LSP) and Speech Data Compression", Proc. of the Int. Conf. Acoust., Speech and Signal Processing, 1.10.1-1.10.4, 1984.

[Jaskie (92)] - C. Jaskie, B. Fette, "A Survey of Low Bit Rate Vocoders", 1992.

[Juang (84)] - B. H. Juang, "On the Hidden Markov Model and dynamic time warping for speech recognition - A unified view", AT&T Tech. J. vol 63. Nº7, pp. 1213-1243, September 1984.

---

<sup>1</sup> Anual, semestral, trimestral, ...

<sup>2</sup> Número total de horas de trabalho.

<sup>3</sup> Discriminadas por tipo de metodologia adotado (T - Ensino teórico; TP - Ensino teórico-prático; PL - Ensino prático e laboratorial; TC - Trabalho de campo; S - Seminário; E - Estágio; OT - Orientação tutorial; O - Outro).

<sup>4</sup> Assinalar sempre que a unidade curricular seja optativa.